

Alinhamento de Textos e Memórias de Tradução em Ambiente de Empresa

Marcos André Vieira Carvalho

Relatório de Estágio de Mestrado em Tradução Área de Especialização em Inglês

Marcos André Vieira Carvalho,
*Alinhamento de Textos e
Memórias de Tradução em
Ambiente de Empresa, 2015*

Abril, 2015

Relatório de Estágio apresentado para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Tradução, realizado sob a orientação científica da Prof.^a Doutora Maria Zulmira Castanheira.

Aos meus pais

AGRADECIMENTOS

Aos meus orientadores, cujas indicações e apoio foram fulcrais para a realização do estágio e do relatório.

À Ana e ao Lein da KennisTranslations SA, pelo seu acolhimento, disponibilidade, bom humor, orientação e ajuda.

Àqueles que fazem parte do meu círculo próximo.

ALINHAMENTO DE TEXTOS E MEMÓRIAS DE TRADUÇÃO EM AMBIENTE DE EMPRESA

MARCOS ANDRÉ VIEIRA CARVALHO

RESUMO

PALAVRAS-CHAVE: tradução, ferramentas TAC, corpora, memória de tradução, alinhamento de texto

Com o constante desenvolvimento das tecnologias para a tradução e a sua importância para o mercado de trabalho, os computadores e as ferramentas TAC tornaram-se instrumentos imprescindíveis ao tradutor. Tendo em conta a crescente preocupação com a otimização do processo de tradução, impõe-se que o tradutor construa as suas próprias fontes de consulta fiáveis, seja a partir da Web, de outros tradutores e até do seu próprio esforço de trabalho prévio.

É possível adquirir memórias de tradução a partir de diversas fontes. Estas memórias têm como principal vantagem o facto de nunca se ter de traduzir o mesmo uma segunda vez, o que proporciona grande economia de tempo e também coerência a nível da terminologia e da fraseologia.

Um dos métodos utilizados na construção de memórias de tradução diz respeito ao alinhamento de textos paralelos, de modo a que os conteúdos desses mesmos textos sejam integrados nas memórias.

Com este relatório pretendeu-se esclarecer algumas questões relacionadas com as duas matérias que figuram no título, abordando as vantagens e desvantagens do alinhamento de documentos para a criação de memórias de tradução e as implicações do seu uso, tanto para o cliente que encomenda a tradução, como para a empresa que trata do projeto e para o próprio tradutor.

TEXT ALIGNMENT AND TRANSLATION MEMORIES WITHIN A COMPANY ENVIRONMENT

MARCOS ANDRÉ VIEIRA CARVALHO

ABSTRACT

KEYWORDS: translation, CAT tools, corpora, translation memory, text alignment

With the constant development of translation technologies and their importance for the labour market, the use of computers and CAT tools has become indispensable to the translator. Bearing in mind the ever growing concern with the optimisation of the translation process, the translator must build his/her own reliable reference sources, either from the Web, other translators and even his/her own prior work effort.

It is possible to acquire translation memories from several sources. These translation memories have as their main advantage the fact that a translator never has to translate the same thing twice, thus providing great savings of time and also consistency in terms of terminology and phraseology.

One of the methods used in the construction of translation memories is the alignment of parallel texts, so that the contents of those texts can be incorporated in the memories.

The aim of this internship report is to clarify some issues related to the two topics mentioned in its title, addressing the advantages and disadvantages of document alignment to create translation memories and the implications of their use, both for the client who commissions the translation, the company that deals with the project and the translator him/herself.

ÍNDICE

Introdução	1
Descrição da empresa	3
1. Enquadramento teórico.....	5
2. Memórias de tradução e alinhamento de textos	11
2.1. Relevância para a tradução	11
2.2. Alinhamento de textos.....	13
2.3. Memórias de tradução.....	18
2.3.1. Criação	18
2.3.2. Avaliação	19
2.3.3. Questões a ponderar	22
3. Estudos de caso	25
Conclusão.....	28
Bibliografia	31
Anexo 1	i

LISTA DE ABREVIATURAS

LC — Língua de Chegada

LP — Língua de Partida

MT — Memória de Tradução

TAC — Tradução Assistida por Computador

TC — Texto de Chegada

TP — Texto de Partida

Introdução

Num mundo profissional caracterizado por crescente competitividade, a qualidade e a eficiência de um tradutor são cada vez mais imprescindíveis. Para que tal aconteça, o tradutor tem à sua disposição várias ferramentas que lhe permitem alcançar esses objetivos, fazendo parte delas as memórias de tradução (doravante referidas por MT). Para que o tradutor evite perder tempo a pesquisar termos repetidos em cada novo trabalho, e para que seja mantida a coerência a nível da terminologia, as MT são uma grande ajuda, pois tornam possível pesquisar e rentabilizar traduções previamente feitas, o que proporciona grande economia de tempo, logo, aumento da produtividade, além de garantirem mais qualidade.

A presente **Introdução** explicita, de forma geral, os objetivos visados no relatório, incluindo também uma apresentação da empresa onde o estágio foi realizado e as tarefas em que mais me concentrei ao longo do mesmo.

No **ponto 1** é feito um enquadramento dos pilares teóricos que sustentam os conhecimentos e aplicações práticas envolvidos no estágio, dando a conhecer visões de autores que se debruçaram sobre os processos de criação de MT e de alinhamento de textos, com maior ênfase naqueles cujas ideias mais se adequam aos objetivos pretendidos com este relatório.

No **ponto 2** é aprofundada a breve contextualização teórica feita anteriormente, sendo dado destaque à importância que as MT têm no desempenho da profissão de tradutor (2.1.). No ponto 2.2. aborda-se uma das duas matérias tematizadas no título do presente relatório: o alinhamento de textos. São dadas a conhecer maneiras de alinhar textos paralelos, indicam-se as razões pelas quais é necessário alinhá-los e que fatores podem dificultar tal alinhamento, e ainda o método para alinhar textos que acabei por preferir adotar durante o estágio. Este assunto volta a ser abordado, em maior pormenor, mais adiante. O ponto 2.3. traz de novo à discussão as MT, agora com enfoque não apenas no processo da sua criação (2.3.1), mas prestando também atenção a sistemas de avaliação das mesmas, ao aspeto económico e às vantagens financeiras que a utilização de MT acarreta (2.3.2). No

ponto 2.3.3. são equacionadas algumas perguntas teóricas sobre as MT que o estágio levantou.

No **ponto 3** são apresentados alguns problemas que ocorreram no processo de alinhamento de textos durante o estágio, bem como as soluções encontradas para os superar, procedendo-se também a uma reflexão sobre essas mesmas dificuldades.

Por último, na **Conclusão** tecem-se algumas considerações finais em relação ao trabalho produzido durante o estágio, passando em revista os pontos principais deste relatório. Tal balanço serve também para sugerir outros caminhos de investigação que no futuro possam vir a ser seguidos, relacionados com as matérias que aqui foram tratadas.

Antes de prosseguir, impõem-se alguns esclarecimentos.

Ao longo deste relatório, as referências a tradutores serão sempre relativas a tradutores trabalhando dentro de uma empresa de tradução, salvo referência em contrário.

Quando se refere a qualidade das traduções, o que está em causa não é avaliar a qualidade da tradução da perspectiva da comparação do texto de chegada (TC) com o texto de partida (TP), como Juliane House e outros teóricos concebem. Assim, a questão da qualidade em tradução nada terá a ver com uma análise do perfil textual do TP em relação com o perfil textual do TC (House 1997), nem com diferentes perspectivas de crítica de tradução, mas sim com as apreciações feitas pelo cliente que encomendou a tradução.

O programa maioritariamente utilizado ao longo do estágio para a atividade de tradução, criação de memórias e alinhamento de textos foi o SDL Trados Studio 2014.¹ Outro programa usado com bastante frequência foi o TStream Editor Studio™, mas apenas para traduzir. Esta ferramenta é uma ferramenta proprietária, cujo acesso às MT é em parte interdito aos tradutores, tendo estes apenas a possibilidade de as

¹ O SDL Trados Studio 2014 contém também outros programas importantes relacionados com a atividade de tradução, nomeadamente o SDL MultiTerm (criação e manutenção de bases de dados) e o SDL Passolo (programa orientado para a localização).

consultar e não de as modificar. Quando o trabalho de tradução termina, o contato com as memórias termina também.

Descrição da empresa

O estágio para a obtenção do grau de Mestre em Tradução foi realizado na conceituada empresa portuguesa KennisTranslations SA, que por sua vez é detida a 100% pela empresa holandesa KennisGuild BV. Na sequência da sua integração no grupo KennisGuild, em 2009, e de acordo com a filosofia do grupo, a empresa é operada por *partners* que gerem as suas próprias contas/clientes, apoiados por uma plataforma informática e estrutura de suporte.

O principal conceito de negócio do grupo KennisGuild consiste na personalização de cada trabalho, ou seja, num contacto próximo entre o cliente e o fornecedor do serviço, com o objetivo de aumentar a satisfação de ambos perante o produto concebido. Faz também parte da filosofia da KennisTranslations fornecer aos tradutores independentes os meios e os recursos necessários ao desenvolvimento de um trabalho de qualidade, para que tal satisfação seja obtida. Entre os principais valores da empresa encontram-se a cooperação, a responsabilidade, o trabalho em equipa e a excelência. É de ressaltar a importância atribuída à comunicação, valor imprescindível para a KennisTranslations: os *partners* são incentivados a comunicar todas as suas observações, dúvidas, opiniões, etc., através de vários meios disponíveis, tanto virtuais como face a face. Através da boa comunicação praticada pela empresa, a qualidade dos serviços prestados e a satisfação dos clientes aumentam substancialmente.

A KennisTranslations SA fornece serviços de tradução nas áreas de especialização dos *partners*, nomeadamente nos domínios da arte, música, medicina e farmácia. Devido ao vasto e diversificado leque de clientes, os *partners* da empresa trabalham com vários tipos de textos e traduzem de uma ou mais línguas, exercendo também as funções de gestores de projeto.

O estágio realizado na KennisTranslations teve a duração de 400 horas, tendo tido início no dia 1 de outubro de 2014 e sido concluído em 1 de Dezembro de 2014.

Dividiu-se em duas partes: uma dedicada à tradução de vários tipos de texto, nomeadamente textos relacionados com arte (arquitetura, pintura, escultura, etc...), em dois pares de línguas (EN>PT e ES>PT), e textos técnicos (manuais de instruções de máquinas de pipocas, textos sobre acessórios de terapias, atas de reunião, manuais de utilizador, entre outros), no par de línguas EN>PT; e uma segunda parte dedicada ao alinhamento de vários tipos de texto e à criação de MT, temas deste relatório de estágio.

1. Enquadramento teórico

Há muito que não se questiona os benefícios da utilização de ferramentas TAC (Tradução Assistida por Computador).² Isto acontece pois a constante evolução destas ferramentas permitiu que o seu uso seja muito mais vantajoso do que a sua não utilização. Antes do século XXI, a dificuldade em utilizar computadores e as próprias ferramentas constituía uns dos problemas associados ao pouco desenvolvimento destas tecnologias, em termos de interface de utilizador e de dificuldades no processamento de textos por parte do *software*. Hoje em dia, praticamente todos os que são, ou aspiram a ser, tradutores a tempo inteiro não questionam se devem manter ou adquirir uma ferramenta TAC. A economia de tempo por elas proporcionada, a utilidade das MT e de outras funcionalidades que facilitam o processo tradutório são fatores decisivos na escolha da aquisição das mesmas. Mas para quem exerce a tradução a título de atividade secundária, ou não se dedica a ela a tempo inteiro, será assim tão necessário adquirir uma ferramenta TAC?

A maior parte das ferramentas TAC são dispendiosas e com o tempo vão ficando desatualizadas, dado o constante desenvolvimento e evolução das tecnologias. Adquiri-las é uma questão que tem de ser devidamente ponderada, para que se recupere o investimento feito. Mais à frente (ponto 2.3.) fornecerei exemplos de algumas situações em que esse retorno do investimento feito na aquisição destas tecnologias pode não existir. Tudo dependerá do tempo dedicado às próprias ferramentas e à atividade de tradução.

O ensaio de Martin Kay “The Proper Place of Men and Machines in Language Translation” (1980) foi um dos primeiros a desenvolver o conceito daquilo a que hoje se chama *translator’s workstation*, ou seja, os programas de tradução assistida por computador. Estes programas, atualmente, têm no seu arsenal de funcionalidades sistemas de MT, *software* de localização, busca por concordância e várias outras

² Palumbo define a Tradução Assistida por Computador (em inglês, CAT), da seguinte forma: “Also called ‘computer-aided translation’, computer-assisted translation, or CAT, is translation carried out, generally at a professional level, with the help of specific computer tools aimed at improving the efficiency of the translation process.” (23)

ferramentas terminológicas, todas elas relevantes para o desempenho da atividade de tradutor. O autor afirma:

If the piece of text to be translated next is anything but entirely straightforward, the translator might start by issuing a command causing the system to display anything in the store that might be relevant to it. This will bring to his attention decisions he made before the actual translation started, statistically significant words and phrases, and a record of anything that had attracted attention when it occurred before. Before going on, he can examine past and future fragments of text that contain similar material. (16).

Desde as décadas de 70 e 80 do século passado que existe o conceito de MT (base de dados eletrónica em que estão armazenados textos traduzidos e respetivos TP, geralmente segmentados em unidades curtas, como a frase), tendo sido Martin Kay um dos primeiros a defini-lo e a conceber a estrutura das ferramentas TAC. O autor refere-se no seu ensaio ao que estas devem conter, nomeadamente um dicionário que permita ao tradutor procurar a palavra ou expressão de que necessita, as próprias memórias que contenham traduções prévias, entre outros programas que facilitem a tarefa do tradutor (Kay 12-18). Enumera também alguns problemas que impedem o verdadeiro sucesso da tradução automática.³

Durante o estágio houve um grande contacto com vários corpora (conjuntos de textos organizados de acordo com determinados princípios, processados por computador e que servem de base a análises de vários tipos), os quais merecem especial atenção neste relatório. Federico Zanettin (178) menciona que os estudos sobre corpora aplicados à tradução envolvem a utilização de dois tipos de corpus, em que textos traduzidos podem ser comparados com os seus TP (corpus paralelo), ou com textos coligidos em outro tipo de corpus selecionado de acordo com critérios semelhantes, por exemplo a nível temático ou a nível de género textual (corpus comparável), na mesma língua ou noutra(s). Estes corpora são utilizados para

³ A Tradução Automática, entendida como a tradução que envolve o emprego de programas de computador, inscreve-se no âmbito da TAC. Embora haja quem estabeleça a distinção entre sistemas de tradução puramente automáticos e sistemas de tradução automática que requerem intervenção humana, tal distinção tem vindo a esbater-se, uma vez que a intervenção humana é praticamente sempre necessária, em maior ou menor grau. Sobre esta matéria, seus limites e perspetivas futuras, ver: Baker e Saldanha 2009.

encontrar regularidades nas traduções, na prática dos tradutores ou nas línguas, e permitem observar, nomeadamente, as decisões tomadas pelos tradutores. A partir da análise dos corpora paralelos é possível discernir as estratégias e as escolhas dos tradutores, ao passo que o estudo dos corpora comparáveis põe em evidência como é que o mesmo conteúdo dos textos é expressado numa ou várias línguas.⁴

Ainda de acordo com Federico Zanettin, as MT são a ferramenta que mais alterou a maneira como a profissão de tradutor é exercida (188). O seu conceito-base reside no facto de se poder reutilizar segmentos (também denominados *translation units*) previamente traduzidos, poupando tempo e aumentando a produtividade e a coerência das traduções (Bowker 2003a: 58). Um segmento pode ser uma frase, um parágrafo ou algo que se assemelhe a partes do texto, nomeadamente títulos ou cabeçalhos (Somers 2003b: 34). As MT são mais eficazes quando chega um novo trabalho e o gestor de projeto se apercebe de que esse trabalho é semelhante a outros anteriormente realizados, tais como versões revistas ou atualizadas de manuais técnicos.⁵ Após o trabalho e a MT serem enviados ao tradutor, basta utilizar a MT e o tradutor traduz o que foi revisto ou acrescentado. Mesmo que não sejam encontradas na memória *exact matches* (correspondências exatas), é possível ainda utilizá-la com proveito, pois podem ser encontradas frases semelhantes, as chamadas correspondências parciais (*fuzzy matches*). Nesta situação, o tradutor pode optar por criar uma tradução a partir daí ou ignorar as correspondências parciais (partes do segmento que correspondem ao TP), criando uma nova e inserindo-a na MT.

Uma funcionalidade bastante importante nas MT é a busca por concordância. Esta função permite procurar palavras ou expressões dentro da memória e o resultado será uma lista extensa, com todos os segmentos em que surgem. Assim, o tradutor pode procurar fraseologia específica, para além de terminologia propriamente dita.

O estágio permitiu ganhar maior consciência sobre a importância das ferramentas TAC e adquirir novos conhecimentos com elas relacionados. Efetivamente, o estágio veio complementar a informação teórica fornecida durante a

⁴ Para mais informações sobre esta matéria, consultar Mona Baker 1993 e 1995.

⁵ As MT são também particularmente úteis no caso da tradução de textos que apresentam vocabulário e estruturas frásicas repetitivos.

componente letiva sobre esta matéria, proporcionando uma visão mais prática sobre a utilidade das MT em ambiente de empresa. São estas que permitem às empresas e, consequentemente, aos tradutores, realizar trabalhos de uma forma mais eficaz e rápida, nunca descurando a qualidade, objetivo máximo de uma tradução. Qualidade, neste contexto, como já foi mencionado, refere-se à boa aceitação ou não da tradução por parte do cliente.

Parece-me importante esclarecer melhor qual é o conceito de qualidade em tradução (em si mesmo um conceito relativo, já que depende de múltiplos fatores, nomeadamente de quem avalia, do parâmetro da avaliação, ou seja, se a tradução é avaliada enquanto processo, produto ou serviço...) que eu aplico neste relatório, distinguindo-o, nomeadamente, do que defende Juliane House, investigadora alemã que muito se tem dedicado a esta questão. Nas suas obras de 1977 (*A Model for Translation Quality Assessment*) e de 1997 (*Translation Quality Assessment: a Model Revisited*), Juliane House propõe um modelo sistemático para avaliação da qualidade da tradução de carácter pragmático, o qual tem como vantagens levar em linha de conta os constrangimentos contextuais, as circunstâncias extralinguísticas, a principal função do TP, o tipo de público-alvo, as normas de uso da LC, entre outros fatores, no sentido de encontrar uma base que permita avaliar o tipo de equivalência entre o TP e o TC. A equivalência entre textos dá-se, nesta abordagem, a nível funcional, pelo que a análise do TP assume um papel determinante para as decisões a tomar pelo tradutor e para o avaliador cujo objetivo é averiguar a qualidade da tradução. Para se fazer uma crítica funcional da tradução é preciso, pois, fazer uma comparação entre o perfil textual do TP e do TC (o qual reflete a sua função), atendendo nomeadamente ao assunto, ao género textual, aos canais de comunicação, aos intervenientes na situação comunicativa, ao modo como se dá tal interação a nível linguístico, analisando, por exemplo, os tipos de frase empregues e as estratégias tradutórias escolhidas.

Dito isto, o que se pretende com este relatório não é analisar o perfil textual dos textos a traduzir e traduzidos, mas sim, tão-só, aceitar o parecer do cliente, ou seja, será este a avaliar quão boa é a qualidade do TC, independentemente dos seus conhecimentos teóricos sobre o conceito de qualidade visto da perspectiva dos Estudos de Tradução.

As ferramentas TAC possuem um programa extremamente importante para a criação de recursos linguísticos: o alinhamento de textos, entendendo-se por tal o estabelecimento de correspondências entre elementos na língua de partida (por exemplo, uma frase ou um parágrafo) e na(s) língua(s) de chegada, correspondências essas que podem depois ser tomadas pelo tradutor como materiais de referência, é uma funcionalidade cuja utilidade prática também se tornou bastante evidente durante a experiência de estágio. Dizem Manning e Schütze:

Text alignment is an almost obligatory first step for making use of multilingual text corpora. [...] Text alignment can also be a useful practical tool for assisting translators. In many situations, such as when dealing with product manuals, documents are regularly revised and then each time translated into various languages. One can reduce the burden on human translators by first aligning the old and revised document to detect changes, then aligning the old document with its translation, and finally splicing in changed sections in the new document into the translation of the old document, so that a translator only has to translate the changed sections. (467)

Estes autores abordam o alinhamento de várias perspectivas, dando a conhecer também as de vários outros autores. Infelizmente, as que são mais focadas dizem respeito ao modo como o alinhamento funciona em termos de algoritmos, fornecendo equações matemáticas para a elaboração de um programa de alinhamento. Trata-se de um assunto demasiado técnico e fora dos meus conhecimentos, pelo que não explorarei esta questão.

Manning e Schütze não fazem a distinção entre bitextos (ou seja, pares de documentos bilinguísticos, que são a tradução uns dos outros) e textos paralelos, preferindo, no entanto, utilizar o primeiro termo. Ao longo deste relatório será utilizado o segundo, por ser mais usado por outros autores aqui também referidos, nomeadamente Lynne Bowker e Harold Somers.

Do excerto acima transcrito podem ser tiradas várias conclusões quanto à importância do alinhamento. Em primeiro lugar, as memórias são o local de armazenamento dos corpora, constituindo um repositório de informação cuja qualidade foi reconhecida e aceite previamente — pelo que são assim consideradas uma fonte fiável.

Em segundo lugar, e indo ao encontro do que diz Federico Zanettin (188-189), o alinhamento é bastante importante, sobretudo no que toca à revisão e atualização de manuais de produtos, podendo mesmo ser feita uma generalização e considerar-se que o alinhamento é vital para a tradução de textos técnicos e científicos. Entre as características do texto técnico e científico encontra-se a coerência a nível de terminologia, isto é, o mesmo item não pode ser traduzido de duas ou mais maneiras diferentes. É importante referir que, ao traduzir textos com carácter técnico e científico, é preferível utilizar a memória com as traduções prévias a efetuar o alinhamento. Caso textos anteriores com as mesmas características tenham dimensões consideráveis, perde-se sempre algum tempo a fazer o alinhamento, sendo esta uma tarefa de último recurso, a realizar só no caso de a MT não ser uma fonte de consulta eficaz.

Por último, o alinhamento serve para aliviar o “fardo” dos tradutores: ao fazer um alinhamento de traduções prévias de documentos semelhantes, o tradutor adquire informação que lhe permite ter apenas de traduzir o que foi revisto ou atualizado no novo documento. Apesar de Manning e Schütze exemplificarem o uso de alinhamento na atualização de manuais técnicos ou de produtos, é possível afirmar que essa utilização também poderá ser aplicada, em princípio, a qualquer outro tipo de texto.

2. Memórias de tradução e alinhamento de textos

2.1. Relevância para a tradução

Apesar do subtítulo ambicioso deste ponto do relatório, o que se pretende provar de seguida é a importância das MT para o tradutor, para a empresa de tradução e para o cliente que encomenda o trabalho.

As MT têm uma importância vital para o tradutor, pois são elas que permitem fazer um trabalho rápido e eficaz, garantindo ao mesmo tempo um controlo da qualidade, uma vez que arquivam traduções que foram previamente validadas e aceites, tanto pela empresa de tradução como pelo cliente que encomendou o trabalho. Ao contrário dos tradutores *freelancer*, cujas MT foram criadas pelos próprios, os tradutores que colaboram com empresas de tradução podem ter acesso a memórias fornecidas pelas mesmas com conteúdos de trabalhos anteriores realizados para o mesmo cliente ou na mesma área, com o objetivo de produzir uma tradução com terminologia coerente e constante e a custos menores. Presumindo que um cliente regular encomenda um trabalho longo, e que o prazo necessário para realizá-lo é mais curto que o ideal, os gestores de projeto poderão ter de dividir o trabalho por dois ou mais tradutores. Neste caso, a(s) memória(s) de tradução tornam-se imprescindíveis, pois permitem a produção de uma tradução coerente, sem disparidades, realizada no prazo estipulado.

Em ambiente de empresa, a utilização das mesmas MT para vários clientes, e não de uma memória para cada cliente, é um aspeto que deve ser devidamente ponderado. É preciso analisar os objetivos dos diversos clientes quanto à terminologia e verificar se os textos a serem traduzidos visam públicos-alvo diferentes. O cliente também poderá informar a empresa de tradução de que quer manter a terminologia anterior, ou prefere antes uma nova tradução (razões para tal podem prender-se com o facto de a terminologia estar desatualizada ou não se desejarem ligações com trabalhos anteriores). Após uma reflexão cuidadosa sobre estes pontos é ainda necessário ponderar qual o benefício que a utilização das mesmas MT para vários clientes pode trazer. Se o objetivo do seu uso for apenas o de consulta de fraseologia e construções gramaticais, então, em princípio, não haverá problema. Contudo, é

sempre preciso ter cuidado enquanto se está a traduzir, respeitando-se os pedidos do cliente para que seja produzido um trabalho de qualidade do qual o tradutor se possa orgulhar e que agrade ao cliente. Na empresa onde foi realizado o estágio verificou-se que ocorriam as duas situações: havia uma ou mais MT para cada cliente e, se um novo cliente encomendasse um trabalho com características semelhantes, as MT seriam novamente utilizadas, todavia sem fazer a sua atualização, sendo o conteúdo da nova tradução inserido numa MT específica para esse novo cliente. Posteriormente, as MT poderiam ser agrupadas por área.

Por exemplo, no SDL Trados Studio 2014, antes de se começar mesmo a traduzir há a opção de apenas utilizar a memória de outros trabalhos como ferramenta de consulta (através da busca por concordância no ato tradutório), sem no entanto atualizá-la com o novo segmento que foi traduzido. Como cada caso é diferente, cabe ao gestor de projeto decidir em que memória(s) colocar as traduções.

Outra questão passível de ser analisada no que toca à utilização de MT diz respeito aos menores custos para a empresa de tradução/tradutor e para o cliente. Uma vez que o esforço de tradução diminui consideravelmente com o uso de uma boa MT, faz sentido que o custo do trabalho diminua proporcionalmente. Para determinar a eficiência que uma MT terá quando for utilizada num trabalho de tradução, é necessário recorrer a outro programa, muitas vezes incluído nas ferramentas TAC, que faz a contagem de palavras. Quantos mais segmentos a MT ajudar a traduzir, maior será a sua eficiência e menor será o custo do trabalho para a empresa, tanto numa perspectiva financeira como do esforço desenvolvido pelo tradutor.

Durante o estágio, foi possível constatar que, durante o processo de gestão de projeto, as MT eram extremamente importantes na criação de um orçamento para o cliente e para o tradutor. Sabe-se que traduzir de uma língua específica tem um custo diferente do que traduzir de outra (por exemplo, traduzir a partir do inglês terá um custo diferente do que traduzir a partir do árabe), assim como fazer alinhamento de textos também terá custos diferentes consoante as línguas envolvidas. Do mesmo modo, o recurso a MT também tem impacto sobre o preço a cobrar ao cliente ou a pagar ao tradutor.

Após o que foi dito, é possível concluir que as MT são vitais para que as traduções sejam realizadas mais depressa. Permitem detetar repetições encontradas nos TP (*exact* e *fuzzy matches*), encontrar terminologia e reutilizá-la em novas traduções, assegurando assim a consistência terminológica, possibilitam também que prazos apertados sejam cumpridos e, ainda, que os custos de determinado projeto sejam mais favoráveis para o cliente. Porém, o que acontece quando precisamos de determinada informação (terminologia, fraseologia, etc...) que já foi traduzida mas não se encontra em nenhuma MT? A resposta a esta pergunta será desenvolvida abaixo.

2.2. Alinhamento de textos

Alignment involves matching up the source text and the translation segment by segment into translation pairs. "Segments" are usually understood to correspond to sentences or other more or less easily distinguishable text portions, such as titles. If the translation is straightforward, then so is the alignment. (Somers 2003b: 34)

O alinhamento é uma funcionalidade que permite a comparação entre o TP e o TC, sendo possível tirar várias conclusões sobre a tradução a partir da sua análise, nomeadamente as opções escolhidas pelo(s) tradutor(es), os problemas e dificuldades de tradução e respetivas soluções.

Harold Somers (2003b) refere que o alinhamento está relacionado com a correspondência entre segmentos do TP e do TC. Estes segmentos, como já foi dito, podem dizer respeito a frases ou frações de frases, e poderão ser consultados futuramente e utilizados como modelos para novas traduções; caso haja uma correspondência igual à tradução que se pretende, basta copiar e colar. Caso os segmentos, no TP e no TC, sejam muito similares em extensão (nomeadamente, com um número aproximado de caracteres e pontuação igual), o alinhamento dos textos será bastante simples e fácil de concretizar.

Grande parte do estágio foi dedicado à criação de MT através do alinhamento de textos paralelos. Há várias razões que levam a que seja necessário criar MT através deste processo, que permite reutilizar o esforço prévio dos tradutores. No final do ponto anterior coloquei uma questão relacionada com um determinado cenário: o

caso em que as traduções são feitas sem recurso a MT, seja porque no momento de tradução não foi utilizada uma ferramenta TAC, ou porque o TP não é processado pela mesma. É o caso de documentos de imagem ou PDF não editáveis, em que a “solução” mais fácil será traduzir num editor de texto com o TP ao lado.

As MT criadas durante o estágio tiveram por base as traduções realizadas por mim e traduções feitas anteriormente por vários tradutores que trabalham com a empresa KennisTranslations. Os textos que foram alinhados têm, na sua génese, várias características que os diferenciam. Grande parte do volume dos alinhamentos efetuados está relacionada com o campo das artes, nomeadamente exposições ao público, e folhetos informativos sobre parques naturais, passando também por traduções de atas e minutas e textos do domínio da economia, entre vários outros tipos de textos. O que se pretende analisar neste relatório não é a dificuldade que cada tipo de texto apresenta ao traduzir, mas sim as complexidades e problemas comuns que todos os tipos de textos e respetivas traduções podem apresentar quando se pretende alinhá-los.

Segundo Harold Somers (2003b: 34), há três fatores que podem dificultar o alinhamento de textos: a dificuldade por parte do *software* em reconhecer onde as frases começam e terminam; uma frase em determinada língua pode não corresponder a uma única frase noutra; o tradutor pode mudar muito ou pouco a ordem das frases na tradução. Estes problemas alertam-nos para as diferenças entre as línguas, as quais muitas vezes obrigam à recolocação de determinados segmentos do texto aquando do processo tradutório, o que cria problemas adicionais para os tradutores que consultam MT como materiais de referência.

O primeiro fator diz respeito às convenções de pontuação das línguas. Nem sempre o programa de alinhamento consegue reconhecer o início e o fim de uma frase, e pode também não reconhecer o título de uma obra ou os títulos dos capítulos de um livro, uma vez que estes normalmente não têm pontuação que indique o seu fim. Assim, quando se for alinhar um texto, o programa irá sempre associar essa “sequência de caracteres” (como Somers lhe chama) ao próximo segmento, juntando dois segmentos que deveriam estar separados.

O segundo e o terceiro fatores dizem respeito às escolhas e decisões do tradutor ao optar por não traduzir determinado texto frase a frase, por exemplo. Bowker (2003b: 96) e Manning e Schütze (467-468) falam também destes fatores. O tradutor pode encurtar, modificar ou prolongar a extensão da frase e/ou dividi-la em duas ou três por questões estilísticas, ou por outras convenções que queira cumprir. Sabendo que o programa de alinhamento assume vários pressupostos e que um deles é o de que a um segmento no TP corresponde outro no TC, ficamos a saber que isto poderá originar um mau alinhamento por parte do programa, caso essa situação não se verifique. Como as línguas apresentam grande variação a nível da ordem de palavras — exemplo do japonês, mandarim, árabe, entre muitas outras, quando comparadas, por exemplo, com as línguas românicas ou germânicas —, a tradução de certas frases a partir de determinadas línguas pode também não resultar no mesmo número de frases no TC.

O processo de alinhamento de textos durante o estágio consistiu nas seguintes fases: receção dos textos a serem alinhados; pré-processamento dos textos paralelos; criação das MT (fase a ser discutida no ponto 2.3.1.); alinhamento automático por parte do *software*; (possível) alinhamento manual; importação do alinhamento para a MT e subsequente verificação; envio das MT para os *partners*.

Os alinhamentos efetuados ao longo do estágio disseram respeito aos corpora enviados por parte dos *partners* da KennisTranslations cujo objetivo era reutilizar essas traduções para projetos futuros. Sendo assim, recebi vários documentos com o objetivo de serem criadas MT a partir deles. Na fase inicial é bastante importante criar um sistema de armazenamento eficaz, especialmente no que toca ao alinhamento de muitos textos, para que não haja confusões ou erros de qualquer tipo.

Após receber os textos passei à segunda fase, que consiste em analisar, em primeiro lugar, os nomes dos documentos e ver se os TP e TC são facilmente reconhecíveis. Se não o forem, basta alterar os nomes dos documentos para que se tornem perfeitamente identificáveis (por exemplo: nome do documento_en.docx; nome do documento_pt.docx). De seguida, faz-se uma leitura dos textos à procura de abreviaturas ou segmentações diferentes nos dois textos que possam constituir problemas ao alinhamento automático. Segundo Tony Hartley, “alignment tools allow

the user to specify how punctuation should be taken into account. Even so, inconsistencies — both ‘UN’ and ‘U.N.’ — can result in wrong segmentation” (112). Se fosse o próprio tradutor a realizar o alinhamento, este saberia quais as abreviaturas mais frequentes no texto e iria incluí-las na lista de exceções do programa de alinhamento para não haver segmentação desnecessária. Neste caso, tal não foi exequível, por ter sido eu a realizar o alinhamento automático. Quando existem vários documentos para serem alinhados, verifica-se um grande gasto de tempo na leitura abrangente e exaustiva de todos os textos, sendo então preferível fazer uma leitura breve dos documentos e proceder mais tarde a um alinhamento manual. Bowker (2003b: 97) refere que, para evitar uma segmentação desnecessária, podem também ser eliminadas as quebras de linha, pois os programas de alinhamento têm tendência para presumir que essas quebras de linha correspondem a mudanças de parágrafos. Assim, a edição dos textos é fulcral na fase de pré-processamento para obter um bom alinhamento.

Após estas duas fases, utilizei o SDL Trados Studio 2014 para fazer o alinhamento automático. Consoante os objetivos pretendidos, é possível alinhar apenas um TP e respetivo TC, ou alinhar vários ficheiros de uma só vez. A escolha de uma destas opções origina diferentes janelas dentro do programa. Contudo, em ambos os casos o próximo passo será sempre criar uma MT, ou utilizar uma já anteriormente criada. Na maior parte das situações foi necessário criar efetivamente uma MT.

A seguir, pode eventualmente ser necessário recorrer ao alinhamento manual, caso o automático tenha dado origem a erros de segmentação.

A penúltima fase consiste na importação dos conteúdos para a MT e na respetiva verificação, para garantir que a importação foi corretamente efetuada. O *software* do SDL Trados Studio 2014 permite fazer a importação dos segmentos com ou sem *tags* (marcas que contêm informação sobre o segmento) e, ao proceder a essa importação, permite igualmente substituir segmentos iguais já existentes na MT (caso se pretenda fazer uma atualização dos conteúdos da MT, ou seja necessário introduzir uma correção na importação feita anteriormente). Apesar de haver outras que permitem fazer uma importação mais personalizada das memórias, as opções de importação do texto com ou sem *tags* e a substituição de segmentos já existentes na

MT foram as opções mais utilizadas durante o estágio para realizar a importação do alinhamento para a MT.

Após estas fases, enviei as MT e os alinhamentos aos *partners* que tinham solicitado a realização de tais trabalhos. Posteriormente, algumas delas foram utilizadas em trabalhos encomendados por clientes.

A questão dos alinhamentos levanta uma pergunta pertinente: quais as consequências de um mau alinhamento? Para responder basta referir que fazer o alinhamento de textos, especialmente os de grandes dimensões, acarreta sempre riscos, pois caso o alinhamento de vários segmentos esteja errado haverá consequentemente uma informação errada que será dada ao tradutor que no futuro venha a consultar a MT. Apesar de se saber que o alinhamento automático nem sempre é perfeito, o alinhamento errado comporta também custos, pois tal implica uma perda de tempo no que toca à verificação e à necessidade de fazer um alinhamento manual dos segmentos errados quando se poderia estar a traduzir.

Manning e Schütze afirmam:

The commonest case of one sentence being translated as one sentence is referred to as a 1:1 sentence alignment. Studies suggest around 90% of alignments are usually of this sort. But sometimes translators break up or join sentences, yielding 1:2 or 2:1, and even 1:3 or 3:1 sentence alignments. (468)

A questão das correspondências é bastante importante relativamente ao alinhamento manual dos textos. Quando o alinhamento automático não corre tão bem quanto o esperado, recorre-se a um alinhamento manual que consiste em “desligar” os segmentos errados e voltar a ligar os segmentos correspondentes corretamente. O sistema de alinhamento manual do SDL Trados Studio 2014 só permite ligar até quatro segmentos, ou seja, quando um segmento no TP corresponde a um, dois ou três segmentos no TC (1:1, 1:2, 1:3), ou vice-versa (1:1, 2:1, 3:1), sendo também possível fazer uma correspondência de 2:2, pelo que, quando as correspondências são maiores, não é possível ligá-los. Durante o processo de alinhamento de textos verificou-se que havia ocasiões em que era necessário ligar quatro segmentos a só um segmento.

Quando tal sucedia, o programa de alinhamento não permitia estabelecer a ligação entre esses segmentos. Para contornar a situação havia duas soluções possíveis: cortar/copiar o texto do segmento a mais e colá-lo noutra segmento que ia ser ligado, ou anular o alinhamento e voltar à fase de pré-processamento de texto, de forma a juntar o que se pretende que fique num único segmento. Cancelar o alinhamento quando já se fez uma parte da verificação (e possível correção manual) de textos de grandes dimensões não parece ser a opção mais viável.

2.3. Memórias de tradução

2.3.1. Criação

Como é do conhecimento geral, uma MT é uma base de dados que serve para armazenar traduções prévias com um único objetivo: o de nunca ter de se traduzir a mesma coisa uma segunda vez.

De acordo com Harold Somers existem três maneiras de construir uma MT: “building it up as you go along, importing it from elsewhere, or creating it from a parallel text” (2003b: 33-34).

A primeira é, segundo Somers, a mais fácil de realizar. À medida que se vai traduzindo, vai-se adicionando cada segmento à base de dados. Porém, no início de carreira, o tradutor *freelancer* não terá traduzido o suficiente para que este método lhe seja útil. É sempre necessário, contudo, começar por algum lado, pelo que o tradutor terá de investir cedo neste método.

A segunda maneira também é simples, uma vez que, utilizando a Internet, é fácil encontrar recursos eletrónicos disponíveis que poderão ser usados numa MT, nomeadamente glossários, bases de dados e até mesmo outras MT. Como Somers refere, há muito que se estabeleceu um formato comum (TMX) a nível da extensão dos ficheiros das MT, a qual permite que estas sejam convertidas para esse formato comum e utilizadas em outras ferramentas TAC.

O terceiro método é o mais importante para este relatório, pois corresponde ao que foi seguido durante o estágio: o alinhamento de textos paralelos com a

subsequente criação de várias MT (uma ou mais para cada cliente, consoante o projeto). Contudo, o alinhamento de textos pode trazer vários problemas, como já foi visto anteriormente (ponto 2.2 do relatório). Fora de um ambiente de empresa, o tradutor *freelancer* pode também recorrer à Internet para encontrar corpora e fazer o respetivo alinhamento.

O processo de criação de uma MT simples, em que esta se encontra vazia e o que se pretende é importar todas as traduções para lá, é bastante fácil. Caso queiramos, é possível personalizá-la. Durante o estágio não criei nenhuma MT personalizada. Às que criei dei-lhes um nome, coloquei-as numa pasta com o nome do cliente a quem a memória iria servir e escolhi as línguas de partida e de chegada. Posteriormente, a memória criada foi incorporada no sistema interno da empresa pelo *partner* que a solicitou, sendo indexada por cliente e área.

Somers (2003b) debruça-se também sobre a função que considera mais importante nos sistemas de MT, a da correspondência, e o sistema qualitativo da mesma, que permite ver quão semelhante ou diferente é a tradução armazenada na memória quando comparada com o TP (o denominado *fuzzy matches score*). Contudo, o objeto de estudo deste relatório é a criação de MT, pelo que não me alongarei sobre a forma como estas qualificam as correspondências.

2.3.2. Avaliação

Outra questão pertinente para a elaboração deste relatório, e em específico para o conteúdo desta secção, é avaliar quão úteis são as MT para o tradutor, para a empresa de tradução e para o cliente. Somers (2003b) fala de três critérios de avaliação sobre os sistemas de MT: a facilidade de utilização; os ganhos na produtividade; o algoritmo de correspondência (41-43). O processo de criação de MT no SDL Trados Studio 2014, como já referi, é bastante fácil, sendo este acompanhado por um assistente de configuração.

De acordo com o autor citado, a produtividade obtida com o uso de MT nem sempre é muito alta. Como foi referido anteriormente, no início da carreira o tradutor *freelancer* não terá traduzido o suficiente para que uma MT lhe seja particularmente

útil, do ponto de vista financeiro. Porém, verifica-se uma diferença de opiniões pois, segundo indicações de vários *partners* da KennisTranslations, o aumento de produtividade verifica-se devido a repetições encontradas no texto logo após a primeira tradução (sendo esta opinião corroborada por Lynne Bowker (2005: 19)).⁶ Estas repetições são apresentadas na ferramenta TAC por causa da MT, sendo apenas necessário verificar se a tradução da MT se ajusta ao pretendido, perdendo-se assim menos tempo com a tradução desses segmentos repetidos.

O terceiro critério prende-se com questões muito técnicas e está fora dos meus conhecimentos e dos objetivos deste relatório, pelo que não me deterei nele.

Aproveitando a reflexão de Somers (2003b) sobre os ganhos na produtividade, retomo a questão que foi colocada na introdução deste relatório sobre o interesse da aquisição, ou não, de ferramentas TAC.

Atualmente, para se traduzir são precisos, no mínimo, dois instrumentos: um computador e uma ferramenta TAC. Para os obter é necessário fazer um investimento, o qual não significa, obrigatoriamente, que se vá arranjar mais trabalho ou que se obtenha vantagem sobre outros tradutores. Significa apenas que, se não houver este investimento, a maior parte das empresas de tradução não oferecerá, provavelmente, trabalho. Assim, a aquisição daqueles dois instrumentos não é, em rigor, tanto um investimento, mas sim um custo inerente à profissão. É um requisito básico.

Quem quer adquirir uma ferramenta TAC precisa de avaliar os seguintes fatores: o tempo que se quer passar a traduzir (A), a quantidade de palavras, em média, que um tradutor traduz por dia (B), os dias úteis num ano (C), os anos de vida económica de uma ferramenta (D), o custo de cada palavra traduzida (E), o custo das correspondências exatas utilizando uma MT (F) e o custo das correspondências parciais recorrendo a uma MT (G). Os valores apresentados em anexo (ver Anexo 1) são meramente demonstrativos e não normativos, não sendo apresentado qualquer tipo de julgamento face aos mesmos.

Se considerarmos, por exemplo, que uma pessoa não quer comprar uma ferramenta TAC e passa 25% do seu tempo a traduzir, e que um tradutor normalmente

⁶ Para mais informações sobre ganhos na produtividade, ver Bowker 2005.

traduz em média 2500 palavras por dia, havendo 258 dias úteis num ano, essa pessoa irá traduzir 161 250 palavras num ano. No final de três anos (tendo por referência que uma ferramenta TAC tem uma vida útil de três anos antes de se tornar obsoleta), terá traduzido 483 750 palavras sem a ferramenta TAC. Se multiplicarmos esse valor pelo preço que se quer cobrar por palavra, nomeadamente 0,05€, obtém-se a quantia de 24 187,50€. Este será o custo do seu trabalho sem a utilização de uma ferramenta de tradução.

A aquisição de uma ferramenta TAC, por seu turno, implica estes e ainda outros fatores, nomeadamente o preço cobrado por correspondências exatas e parciais (*exact* e *fuzzy matches*) e a frequência, em média, com que estas ocorrem. Se considerarmos que a mesma pessoa traduz à mesma 483 750 palavras ao ano e que 85% são palavras que nunca foram traduzidas, essa pessoa ganhará 20 559,38€. Pressupondo que as correspondências exatas surgem com uma frequência de 5% do total de palavras traduzido e que o valor cobrado é de 25% do valor cobrado por uma palavra nova (portanto, 0,0125€), o valor ganho será de 302,34€. Presumindo que as correspondências parciais surgem com uma frequência de 10% e que o valor cobrado é 60% do valor cobrado por uma palavra nova (ou seja, 0,03€), o valor será de 1 451,25€. Somando estes três montantes obtemos o preço que determinada pessoa iria receber caso adquirisse uma ferramenta de tradução: 22 312,97€. Acha-se assim uma diferença de 1 874,53€. Apesar desta diferença, é preciso salientar que não se está a perder dinheiro com a utilização de uma ferramenta TAC, mas sim a ganhar-se tempo que será utilizado a traduzir mais, fazendo com que ao fim de três anos o volume de palavras traduzido seja muito mais alto e invertendo-se assim completamente os valores finais. Verifica-se também uma redução de custos, pois há uma diminuição do esforço tradutório.

Relembro que os dados apresentados não são indicadores do mercado de tradução e que a tabela em anexo pretende apenas ajudar a perceber melhor os cálculos realizados para chegar aos valores apontados.

Adicionalmente, há ainda outros fatores que não foram explorados aqui, mas que salientariam as vantagens da aquisição de ferramentas TAC. Por exemplo, há

clientes que insistem na produção de certos trabalhos com recurso às mesmas, o que implica que o tradutor perca o trabalho caso não as tenha.

2.3.3. Questões a ponderar

As MT já deram amplas provas da sua eficácia e qualidade no auxílio ao trabalho do tradutor *freelancer* e do tradutor inserido numa empresa. Contudo, a utilização de MT suscita um conjunto de questões que muitas vezes passam despercebidas. Uma MT é tanto mais útil quanto maior for? Valerá a pena fazer alinhamento de textos com um número reduzido de palavras, ou é preferível utilizar textos com um número elevado de palavras? Há um número mínimo de palavras aceitável?

Estas são algumas das perguntas que me ocorreram logo de início e que pensei que poderiam ser equacionadas neste relatório. Com o decorrer do estágio foi-se verificando uma mudança de perspetiva sobre as mesmas. O primeiro instinto é responder “sim” à primeira interrogação. E talvez seja a resposta certa. Porém, considero que a qualidade de uma MT terá sempre um peso maior do que a quantidade de segmentos nela registada, ou seja, mesmo que uma MT seja menor em número de segmentos quando comparada com outra maior, a primeira poderá ser mais relevante caso contenha a informação desejada para o trabalho em que está a ser utilizada. É sempre possível argumentar que se uma MT for maior, a sua utilidade será também provavelmente maior. Todavia, como foi referido acima, nada garante que a informação pretendida se encontre nessa MT maior. Por exemplo, no caso de uma MT com 300 segmentos referente a um projeto de alinhamento realizado durante o estágio com base em traduções de diversos tipos de certificados, a sua relevância será muito maior para um tradutor que traduza regularmente certificados de natureza variada do que uma MT mais geral, que englobe diversos tipos de documentos (pressupondo que um tradutor tem uma MT geral e outras mais específicas). Tudo se resume à informação contida na memória e não ao seu tamanho.

A segunda questão passível de ser analisada também me suscitou grande interesse durante o estágio. Deverá o tradutor aplicar os seus esforços no alinhamento

de textos com um número reduzido de palavras, ou não compensará o tempo que se leva a pré-processar os textos e a alinhá-los manualmente, caso necessário? Tendo em conta a experiência obtida através do estágio, posso afirmar que só em casos muito específicos é que não valerá a pena fazer os alinhamentos, independentemente da extensão dos textos. Apesar de não ter traduzido textos de carácter literário, o conhecimento adquirido ao longo da componente letiva permite-me chegar a algumas conclusões sobre a pergunta formulada. Em textos técnicos, de maior ou menor dimensão, a probabilidade de haver várias expressões e termos repetidos é muito grande, pelo que o número de palavras não deverá afetar a escolha de fazer o alinhamento ou não, enquanto na tradução de textos literários, ou até determinados textos pragmáticos, não parece haver necessidade de criar uma MT, uma vez que os objetivos da tradução são outros e as escolhas do tradutor são diferentes (certamente que não serão utilizadas traduções prévias em novos poemas, em anúncios ou outros tipos de textos em que a originalidade prevalece e em que a repetição de termos traduzidos pode até, em certos casos, ser considerada uma deficiência de composição, um sinal de pobreza vocabular e estilística). Bowker e Barlow (10) referem que os passos inerentes ao pré-processamento de textos podem ser morosos e que uma verificação manual por parte de quem alinha os textos é sempre necessária, de forma a garantir que o alinhamento foi bem efetuado; apesar de o processo ser demorado, parece-me que é sempre vantajoso fazer o alinhamento dos textos paralelos, independentemente da sua extensão, uma vez que estes poderão ser reutilizados novamente no futuro, a menos que os documentos criem tais dificuldades que seja muito difícil alinhá-los, como será visto adiante.

Relativamente à última questão, não me parece ser necessário estipular um número mínimo de palavras aceitável para que se possa realizar alinhamentos. Houve ao longo do estágio uma mudança de perspetiva que levou a uma nova formulação da pergunta. O que se deve questionar quando se está prestes a fazer alinhamentos é: será que vão aparecer novos trabalhos na mesma área e com terminologia semelhante, ou este é só um trabalho isolado? Se se souber já que no futuro irão aparecer trabalhos semelhantes, a resposta é “sim, sem dúvida que valerá a pena fazer

alinhamentos”.⁷ Se a resposta for uma incógnita e houver tempo livre, por que não realizá-los? Não se perde nada em fazê-lo. Quem sabe se no futuro não serão necessários novamente, não se perdendo nessa altura tempo com tal tarefa e ficando o tradutor mais livre para se poder dedicar à tradução. Durante o estágio recebi vários documentos pequenos (não chegavam a 500 palavras cada, na sua maioria) que não constavam em MT. Apesar de terem uma dimensão modesta, sentiu-se a necessidade de fazer um alinhamento desses documentos, devido precisamente ao facto de já serem vários com a mesma terminologia e fraseologia.

⁷ À medida que o tempo passa, o tradutor vai-se especializando cada vez mais em determinadas áreas.

3. Estudos de caso

Durante o estágio ocorreram várias situações relativamente ao alinhamento de textos que dificultaram essa tarefa. De seguida, fornecem-se alguns exemplos de problemas encontrados. O mais comum foi a diferença entre o número de segmentos do TP e do TC. Isto acontece devido às diferentes convenções de pontuação das línguas de trabalho. Uma vez que a segmentação do texto é feita, principalmente, através da pontuação, isto origina diferentes segmentos nos dois textos. Sabendo *a priori* que o alinhamento de textos dará lugar a este problema, facilmente se consegue resolver a questão. A ferramenta de alinhamento de textos do SDL Trados Studio 2014 contém uma funcionalidade que permite analisar essas pontuações, indicar quão relevante cada pontuação é e como devem ser feitas as segmentações do texto. Como os alinhamentos realizados durante o estágio foram feitos a partir de trabalhos de outros tradutores, é difícil saber se, e como, este tipo de exemplos ocorreu, pelo que foi necessário fazer uma análise dos textos à procura de segmentações devidas a pontuações diferentes.

Ao longo do estágio, como ficou dito acima, foi necessário realizar o alinhamento de vários documentos que foram traduzidos por diversos *partners* devido à extensão de um determinado projeto que a KennisTranslations teve em mãos. Ao todo, nesse projeto, foram traduzidos mais de 100 documentos cuja LP era o espanhol e as LC eram o inglês e o português (totalizando mais de 350 ficheiros). Tal projeto de enorme envergadura consistiu na tradução de textos de carácter técnico, cujos temas estavam ligados à arte antiga (arquitetura, pintura, escultura, etc...), para uma instituição conceituada⁸. Uma vez que as convenções do texto técnico são diferentes das do texto literário e implicam uma maior proximidade entre TP e TC, as traduções deste projeto não sofreram grandes alterações relativamente à segmentação do TP, sendo fácil para o *software* fazer um alinhamento rápido e eficaz, sem grandes desvios. Tratando-se de um corpus bastante extenso, cuja informação era muito específica, verificou-se a necessidade de criar uma MT para cada par de línguas (ES>EN e ES>PT, tendo-me sido pedido que criasse ainda uma MT PT>EN), apesar de não ser um cliente

⁸ Por questões de confidencialidade, não me é possível revelar o nome do cliente.

novo e já existir uma MT para o mesmo. Felizmente, os documentos recebidos eram de fácil identificação e encontravam-se organizados, pelo que não foi necessário verificar quais as línguas de cada documento e quais os seus correspondentes nas outras línguas. Uma vez que participei nesse projeto de tradução, eram do meu conhecimento algumas das abreviaturas mais utilizadas, portanto a inserção das mesmas na lista de exceções do *software* de alinhamento foi de simples execução. Apesar de alinhar todos os documentos ser uma tarefa morosa, foi sempre fácil fazê-lo e poucas foram as vezes em que se tornou necessário recorrer ao alinhamento manual dos textos. Contudo, foi sempre feita uma leitura comparativa do TP e TC para confirmar que o alinhamento automático tinha sido bem-sucedido e que não era necessário fazer alterações manuais. Devido à grande quantidade de documentos envolvida no projeto, sucedeu uma ou outra abreviatura ficar esquecida, contudo, uma vez que este erro ocorria nos dois textos, o número de segmentos continuava igual, não dificultando o alinhamento manual.

Neste estudo de caso, o importante a ser retirado como conclusão é que uma boa organização no sistema de arquivação aquando da receção de documentos é fundamental para agilizar o processo de alinhamento. Sejam muitos ou poucos, é sempre necessário que os documentos sejam identificados com um título apropriado, e não com uma combinação de números e letras, como se verificou em outros documentos recebidos para alinhar, e com a identificação da língua em que foram criados. No meu caso, isto bastou para que a tarefa de alinhar se tornasse muito mais eficiente e rápida, libertando-me tempo para que me pudesse dedicar a outros alinhamentos ou traduções.

Como já foi referido, o estágio proporcionou a oportunidade de fazer vários alinhamentos com vários tipos de documentos e de textos. Na maior parte dos casos, o tipo de documentos tinha uma extensão .docx, formato facilmente reconhecível pelo SDL Trados Studio 2014, tanto para traduzir como para alinhar. Contudo, houve um projeto que consistiu no alinhamento de vários tipos de certificados (certificados de habilitação, de aptidão profissional, de formação profissional, entre outros) cuja extensão de documento era diferente. Foram recebidos documentos PDF não-editáveis e até ficheiros de imagens de certificados com baixa qualidade. Rapidamente

se sentiu a necessidade de adquirir um *software* que permitisse ler estes PDF não editáveis, uma vez que o SDL Trados Studio 2014, apesar de ter um conversor de documentos e um programa de OCR muito bons no seu leque de programas, não era “forte” o suficiente para converter e ler corretamente o texto todo. Utilizando inicialmente programas gratuitos encontrados na Internet para converter os ficheiros, depressa se verificou que a conversão não era bem executada e que o *software* não conseguia “ler” corretamente os ficheiros devido à baixa qualidade das imagens. Quando já reinava algum desânimo, experimentou-se a versão de teste do Adobe Acrobat™. Com este programa rapidamente se começou a conseguir converter alguns ficheiros PDF em documentos do Word iguais ao TP. A partir daí foi fácil inserir os TP e TC. Contudo, os problemas encontrados por mim, enquanto estagiário, tinham sido enfrentados também pelo tradutor. Aquando da tradução, verificou-se que, em vez de serem utilizadas ferramentas TAC, recorreu-se a um processador de texto, nomeadamente o Microsoft Word. Tal opção deveu-se ao facto de, no momento de inserção do documento na ferramenta TAC, esta não reconhecer o documento, ou a conversão ser mal executada, tornando-se necessário então usar o processador de texto. Daí derivou a necessidade de se criar uma MT em que constassem essas traduções.

Este projeto foi o mais complicado de alinhar, pois a fase de pré-processamento implicou várias horas a tentar que o TP e o TC ficassem semelhantes em termos de segmentos. O que se queria evitar era ter de passar manualmente o TP para um processador de texto para que este ficasse, se não igual, pelo menos semelhante. Foi o que se fez com alguns documentos demasiado ilegíveis para que os programas de OCR conseguissem identificar o texto.

É preciso salientar que o processo de alinhamento de textos referido aqui está relacionado com o sistema de trabalho aprendido e desenvolvido a partir da experiência obtida durante o estágio, com vista a satisfazer determinadas necessidades e a criar um método de trabalho organizado e eficiente.

Conclusão

A partir do que foi referido ao longo do relatório tecerei algumas considerações finais sobre o mesmo. Em primeiro lugar, é necessário frisar que a crescente competitividade no mercado de trabalho leva a que seja necessário, e até exigido ao tradutor, o uso de um computador e de uma ferramenta TAC. Os tempos mudaram, não sendo já possível traduzir apenas com papel e caneta e com um dicionário ao lado.

Desde 1980 que foram desenvolvidas as estruturas das ferramentas TAC, tornando-se estas imprescindíveis para um trabalho rápido, eficiente e de elevada qualidade. Uma das vantagens das ferramentas TAC reside nas MT, as quais constituem uma fonte de consulta fiável em termos de terminologia específica e fraseologia. A utilização de MT suscita várias questões. Em primeiro lugar, vimos que usar as mesmas MT para vários clientes deve ser um ato ponderado e com objetivos bem definidos, para que os pedidos do cliente sejam respeitados e os conteúdos não sejam incorretamente misturados. Mostrou-se o impacto das MT para o cliente, para a empresa de tradução e para o tradutor em termos de orçamento e também de esforço de tradução.

Como foi observado neste relatório, há várias maneiras de criar uma MT. No contexto do estágio realizado, foi salientada a que se revelou mais pertinente: o alinhamento de textos paralelos. O alinhamento serve para que se possa integrar traduções previamente validadas e aceites nos recursos linguísticos a serem consultados pelo tradutor. Esta funcionalidade consiste na ligação de um ou mais segmentos no TP ao(s) segmento(s) correspondente(s) no TC, criando assim *translation units* que irão ser guardadas na MT. Contudo, o alinhamento de textos pode ser uma tarefa bastante morosa e entediante (especialmente no que toca a documentos extensos), acarretando alguns riscos e podendo até ser prejudicial caso o alinhamento seja mal feito.

Assim, considero que o alinhamento de textos deve ser utilizado pelo tradutor como último recurso, sendo preferível o uso de MT por questões de tempo e por causa dos riscos referidos no relatório. Contudo, não me refiro aqui aos alinhamentos opcionais feitos quando há tempo disponível para tal, pois há várias vantagens que daí

derivam, como demonstrei atrás. Há maior probabilidade de os erros ocorrerem quando há urgência, pagando-se assim os custos já apontados, nomeadamente os de se perder mais tempo a corrigir o que está errado, e, caso o erro passe despercebido, o de um tradutor vir, no futuro, a utilizar uma informação incorreta.

Um dos conhecimentos mais importantes que adquiri com o estágio diz respeito aos métodos de realizar alinhamentos. No início pouco sabia sobre o assunto, mas fui adquirindo competências e aprendendo formas de resolver os problemas que iam surgindo. Relembro, de forma sucinta, o método adquirido, por me parecer fulcral: criar bons documentos de partida de modo a que a ferramenta de tradução consiga reconhecer o ficheiro e fazer uma boa segmentação (fase de pré-processamento dos textos); criar uma MT com um nome de fácil reconhecimento, seguindo as indicações do assistente de configuração; inserir os documentos no programa de alinhamento; possível alinhamento manual, caso se verifiquem discrepâncias de informação nos segmentos ligados; importar os segmentos (*translation units*) para a MT e proceder a subsequente verificação.

Por último, gostaria de deixar aqui algumas sugestões para futuros trabalhos. Por questões de espaço, tendo em conta o limite apertado de páginas estipulado para um relatório de estágio, não me foi possível abordar todas as questões que este me suscitou, nem me alongar sobre certos assuntos. A busca por concordância parece-me ser um bom tópico de investigação, pois é uma das características que tornam as MT bastante vantajosas. Outro possível tema que descobri durante as minhas pesquisas prende-se com a existência de uma diferença entre *bilingual concordancers* e *translation memory systems*. Tanto quanto sei, os primeiros já não são tão utilizados como os segundos, pelo que me parece que as razões para que tal tenha acontecido mereceriam ser objeto de estudo. Lynne Bowker e Michael Barlow (2008) fazem uma análise comparativa sobre o assunto, sendo assim um bom ponto de partida. Por fim, gostaria de sugerir um estudo sobre a importância das MT para a tradução de textos literários e pragmáticos, uma vez que este relatório se concentrou essencialmente na relevância daquelas para a tradução de textos técnicos. Devido à universalidade do uso de ferramentas TAC na tradução, parece-me pertinente investigar o interesse das MT

para a tradução de textos de carácter não técnico, para verificar se também nesse caso trazem vantagens, ou se, pelo contrário, não têm aplicabilidade.

Por último, gostaria de referir que, apesar de não dar nenhum exemplo em concreto de traduções ou alinhamentos feitos por mim durante o estágio, devido à regra da confidencialidade, procurei equacionar as matérias abordadas a partir de exemplos gerais que fossem claros e ilustrativos.

Bibliografia

BAKER, Mona. "Corpora in Translation Studies: An Overview and Some Suggestion for Future Research". *Target* 7 (2), 1995: 223-43.

BAKER, Mona and Gabriela Saldanha, eds. *Routledge Encyclopedia of Translation Studies*. Second edition. London and New York: Routledge, 2009. (1st ed. 1998).

BAKER, Mona, G. Francis and Elena Tognini-Bonelli, eds. *Text and Technology: in Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 1993.

BENIS, Michael. "Translation Memory from O to R". *ITI Bulletin*, April 1999: n. pag.
Acedido em: 13 Março 2015: <http://utkl.ff.cuni.cz/~rosen/VYUKA/MT/tm-review01.htm>.

BOWKER, Lynne. "Productivity vs Quality? A Pilot Study on the Impact of Translation Memory Systems". *Localisation Focus – The International Journal of Localisation*, 4-1, March 2005: 13-20. Acedido em: 10 Abril 2015: http://www.localisation.ie/sites/default/files/publications/Vol4_1Bowker.pdf

---. "Terminology Tools for Translators". *Computers and Translation: a Translator's Guide*. Edited by Harold Somers. Amsterdam and Philadelphia: John Benjamins, 2003b. 49-66.

BOWKER, Lynne and Jennifer Pearson. *Working with Specialized Language: A Practical Guide to Using Corpora*. London and New York: Routledge/ Taylor & Francis Group, 2003a.

BOWKER, Lynne and Michael Barlow. "A Comparative Evaluation of Bilingual Concordancers and Translation Memory Systems". *Topics in Language Resources for Translation and Localisation*. Edited by Elia Yuste Rodrigo. Amsterdam and Philadelphia: John Benjamins, 2008. 1-22.

HARRIS, Brian. "Bi-text, a new concept in translation theory". *Language Monthly* 54.8-10, March 1988: n. pag.. Acedido em: 19 Março 2015: <http://mt-archive.info/LangMonthly-54-1988-Harris.pdf>.

HARTLEY, Tony. "Technology and Translation". *The Routledge Companion to Translation Studies*. Edited by Jeremy Munday. London and New York: Routledge, 2009. 106-127.

HOUSE, Juliane. *A Model for Translation Quality Assessment*. Tübingen: Gunter Narr, 1977.

---, ed. *Translation: A Multidisciplinary Approach*. Basingstoke, UK and New York: Palgrave Macmillan, 2014.

---. *Translation Quality Assessment: A Model Revisited*. Tübingen: Gunter Narr, 1997.

KAY, Martin. "The Proper Place of Men and Machines in Language Translation". *Research Report CSL-80-11*. Palo Alto, California: Xerox PARC, 1980. Reprinted in *Machine Translation* 12 (1997). 3-23.

MANNING, Christopher D. and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. Cambridge, Massachusetts: MIT Press, 1999.

McKAY, Corinne. *How to Succeed as a Freelance Translator*. Second Edition. New York: Two Rat Press, 2011.

MELBY, Alan. "Multi-level Translation Aids in a Distributed System". *Proceedings of the Ninth International Conference on Computational Linguistics*. Edited by J. Horecky. Praga: Academia, 1982.

MILLÁN, Carmen and Francesca Bartrina, eds. *The Routledge Handbook of Translation Studies*. London and New York: Routledge: 2013.

PALUMBO, Giuseppe. *Key Terms in Translation Studies*. London and New York: Continuum, 2009.

RODRIGO, Elia Yuste. *Topics in language resources for translation and localisation*. Amsterdam and Philadelphia: John Benjamins, 2008.

SOMERS, Harold. *Computers and Translation: a Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, 2003a.

---. "Translation Memory Systems". *Computers and Translation: a Translator's Guide*. Edited by Harold Somers. Amsterdam and Philadelphia: John Benjamins, 2003b. 33-47.

ZANETTIN, Federico. "Corpora in Translation". *Translation: a Multidisciplinary Approach*. Edited by Juliane House. Basingstoke, UK and New York: Palgrave Macmillan, 2014. 178-199.

Anexo 1

Pressupostos:

Ocupação de tempo — 25 % (A)

Quantidade de palavras que um tradutor traduz, em média, por dia — 2500 (B)

Dias úteis num ano — 258 (C)

Vida económica de uma ferramenta TAC em anos — 3 (D)

Custo por palavra — 0,05 € (E)

Custo das correspondências exatas — 0,0125 € (F)

Custo das correspondências parciais — 0,03 € (G)

Frequência, em média, de correspondências exatas — 5% (H)

Frequência, em média, de correspondências parciais — 10% (I)

Percentagem de custo sobre o valor total do custo por palavra relativo às correspondências exatas — 25% (J)

Percentagem de custo sobre o valor total do custo por palavra relativo às correspondências parciais — 60% (K)

Fórmula para cálculo sem ferramenta:

$$A \times B \times C \times D \times E = 24\,187,5\text{€}$$

Fórmula para cálculo com ferramenta:

$$(A \times B \times C \times D \times E \times (1 - J - K)) + (A \times B \times C \times D \times F \times H) + (A \times B \times C \times D \times G \times I) = 22\,312,97\text{€}$$